|  |  |
| --- | --- |
| **ITU – Telecommunications Standardization Sector**STUDY GROUP 16 Question 6**Video Coding Experts Group (VCEG)**73rd Meeting: 17-26 January 2024, Teleconference | Document VCEG-BU01 |

|  |  |
| --- | --- |
| Question: | Q.6/SG16 (VCEG) |
| Source: | **Christian Helmrich, Christian Rudat, Jonathan Pfaff, Heiko Schwarz, Detlev Marpe, Thomas Wiegand (Fraunhofer HHI)** | Tel: Email: | firstname.lastname@hhi.fraunhofer.de |
| Title: | **Performance evaluation of audio codecs for multichannel biomedical data** |
| Purpose: | Information |

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

# Introduction

This contribution reports on a continuation of the investigation described inVCEG-BT05 in order to extend the scope towards the identification of a reference codec for compressing biomedical waveform data with more than two channels. It reports additional compression results for the Extended HE-AAC audio codec, when operated without psychoacoustic encoder optimizations and encoding each channel individually, on the Ozdemir EMG dataset (4 channels per sequence) and the MUSC EEG set (up to 276 channels per sequence). The performances of Extended HE-AAC with and without a restriction to using only short-transform encoding are compared as well.

# Multichannel encoding with MPEG-D Extended HE-AAC

Since all publicly available Extended HE-AAC decoder implementations are restricted to at most two channels [1], it was decided to perform the reference encoding of biomedical data with more than two channels using this audio codec by encoding each channel individually. To this end, the individual files of the two multichannel datasets listed in VCEG-BT07 [2], the EMG “Ozdemir” and the EEG “MUSC” dataset, were repackaged (without alteration of the PCM sample data) into single-channel 32-kHz Wave (.wav) formatted audio files, which were then encoded separately using the exhale encoder [3] and the same configuration of encoding parameters and source code modifications as described in VCEG-BT05 [4]. In particular, bit rate preset # ranged from 4 to 8.

Since # allows for only relatively coarse variation of the bit rate and at least eight BPS vs. PRD points were desired to cover the full range of lossy working points (WP1–WP8) as much as pos­sible, intermediate operating points were generated by changing in line 972 of lib/exhaleEnc.cpp,

 s = \_\_max (1u + ((UINT32\_MAX / (eightShorts … (1)

as follows, thereby effectively halving the quantizer step size:

 s = \_\_max (1u + ((INT32\_MAX / (eightShorts … (2)

On the FTP server, the bitstreams are distinguished by *uint* and *int*, respectively, in their names. Moreover, the additional bugfix commit1ba38din exhale’s upstream Git repository was included.

# Encoder restriction to the usage of only short transforms

As requested at the 73rd VCEG meeting in Hannover, a comparative evaluation was performed between two variants of the Extended HE-AAC encoder employed in this investigation, namely,

* exhale without psychoacoustic encoder optimizations and with per-frame input dependent transform length adaptation (also called “block switching” in MPEG audio terminology),
* exhale without psychoacoustic encoder optimizations and with encoding restricted to the use of only short transforms (“eight short blocks” in MPEG terminology) in every frame.

Figures 1 and 2 illustrate that, for each of the three datasets listed in VCEG-BT07, the transform length adaptive encoder outperforms the short-block constrained encoder by a significant margin and that the choice of PRD calculation has almost no effect on the relative locations of the lines.

Figure 1 also shows that, on average across the dataset, joint-stereo coding of the ECG sequences provides, due to the lack of more sophisticated MSE friendly encoder optimizations, almost no benefit over independent mono (i.e., single-channel) encoding of each data channel. Therefore, it is suggested to, for consistency and harmonization reasons, employ the single-channel encodings as audio reference conditions also for the MIT set in future comparative evaluation experiments.


bits/sample

**–** exhale stereo (VCEG-BT05)

**–** exhale mono

**–** exhale mono shortonly

**–** exhale stereo (VCEG-BT05)

**–** exhale mono

**–** exhale mono shortonly

**Figure1**. Performance comparison on ECG set for different encoder settings with (*left*) PRD cal­culation used in VCEG-BT05, (*right*) variance based PRD calculation described in VCEG-BT07.


bits/sample

**–** exhale mono

**–** exhale mono shortonly

**–** exhale mono

**–** exhale mono shortonly

**Figure 2**. Performance comparison on (*left*) EMG and (*right*) EEG dataset for different encoder settings and the PRD calculation used in VCEG-BT05. Only every second bit rate point is tested on the two datasets in order to reduce the computational burden (the sets contain many channels).

# Experimental results

Figures 3, 4, and 5 illustrate the performance curves resulting from the experiments described in the previous sections, averaged arithmetically across the performance datapoints for the individual sequences of each dataset. In each figure, the left graph depicts the results for variance normalized percentage root mean square distortion (PRD) as defined in sec. 4 of VCEG-BT07 (called PRD\_var in the figures), while the right graph depicts the results for the variance and channel normalized percentage root mean square distortion (CPRD), also specified in sec. 4 of VCEG-BT07. It must be noted that some channels were found to exhibit zero overall input variance. Since such variance values prevent the calculation of meaningful CPRD averages (divisions by zero), PRD values of zero were assumed for the affected channels. This assumption is considered reasonable because, in all but one of the few affected channels, exhale encoded the (constant) PCM values losslessly.

Per-sequence graphs are provided on the FTP server (see sec. 5) in zip file *VCEG-BT08\_graphs.zip*.

 

**Figure 3**. Overall mean encoding performance for ECG set in terms of (*left*) PRD, (*right*) CPRD.

 

**Figure 4**.Overall mean encoding performance for EMG set in terms of (*left*) PRD,(*right*) CPRD.

 

**Figure 5**. Overall mean encoding performance for EEG set in terms of (*left*) PRD, (*right*) CPRD.

# Availability of data and software

The test data (encoded audio bitstreams, PDFs with graphs, and CSV tables) as well as the software modifications applied to generate the above results for Extended HE-AAC can be downloaded from the following location, where they are stored as files or directories beginning with *VCEG-BT08\_*:

Server: [ftp.hhi.fraunhofer.de](https://urldefense.com/v3/__http%3A/ftp.hhi.fraunhofer.de__;!!Ab1_Rw!ExOfc-x3p0e9RCmWEdTCbSatvCU5AIjYpC7ovhXAk9GmDH_epQgMsHc6UnSLIewoZqsap0lXQl4o7Z98T51cvaymJphwNcozfzg$)

Login: dicom

Password: yX5GUw.Zn

# References

1. GitHub, “Cannot decode USAC with greater than 2 channels, or possibly dual mono,” 2020. <https://github.com/mstorsjo/fdk-aac/issues/120>.
2. J. Pfaff and J. Halford, “Call for Evidence on the coding of biomedical waveform data,” Q.6/SG16, *doc.VCEG-BT07*, Hannover, Nov. 2023.
3. C.R.Helmrich, project ecodis, “exhale: ecodis extended high-efficiency and low-complexity encoder,”version1.2, *Gitlab repository*, Nov.2023. <https://gitlab.com/ecodis/exhale>. Latest commit: <https://gitlab.com/ecodis/exhale/-/commit/1ba38d589>.
4. C.R.Helmrich *et al.*, “Inform. on performance evaluation of audio codecs for 2-channel ECG data,” *doc. VCEG-BT05*, Hannover, Oct. 2023.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_